NAVAL POSTGRADUATE SCHOOL
Monterey, California

Methodology for an Operationally-Based Test Length Decision

by

Donald P. Gaver
Patricia A. Jacobs

August 1997

19970918 134

Prepared for: Defense Operational Test and Evaluation
The Pentagon, (Room 3E318)
Washington, DC 20301-1700

DTIC QUALITY INSPECTED 3

NAVAL POSTGRADUATE SCHOOL
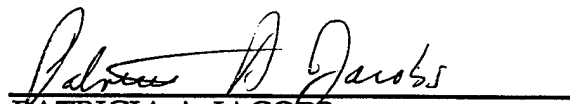**MONTEREY, CA 93943-5000**

Rear Admiral M. J. Evans
Superintendent

Richard Elster
Provost

This report was prepared by:

DONALD P. GAVER
Distinguished Professor of
Operations Research
Naval Postgraduate School

PATRICIA A. JACOBS
Professor of Operations Research
Naval Postgraduate School

Reviewed by:

Released by:

RICHARD E. ROSENTHAL
Chairman
Department of Operations Research

DAVID W. NETZER
Associate Provost and Dean of Research

GERALD G. BROWN
Associate Chairman for Research
Department of Operations Research

REPORT DOCUMENTATION PAGE

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE August 1997	3. REPORT TYPE AND DATES COVERED Technical

4. TITLE AND SUBTITLE

Methodology for an Operationally-Based Test Length Decision

5. FUNDING NUMBERS

DVAM70017

6. AUTHOR(S)

Donald P. Gaver and Patricia A. Jacobs

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Naval Postgraduate School
Monterey, CA 93943

8. PERFORMING ORGANIZATION REPORT NUMBER

NPS-OR-97-015

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Defense Operational Test and Evaluation
The Pentagon (Room 3E318)
Washington, DC 20301-1700

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

Weapon systems that function destructively (e.g. missile or torpedo) are to be acquired in a lot of size m. Acceptance of the lot is based on the result of an operational test, administered to part of the lot: if the test results indicate positive operational value the lot is accepted and the remaining part of the lot if fielded; otherwise the lot is "rejected".

A test plan is designed that establishes an optimal number of weapon copies to test, given (models of) the operational gain of the fielded weapon under two tactical options, and the uncertainty in the weapon's predicted probability of success after the test is complete. The major test objective is to realize possible operational utility from the lot of items, and secondarily to demonstrate arbitrary levels of certainty.

14. SUBJECT TERMS

Test and Evaluation, combining information, meta-analysis, Bayesian Models

15. NUMBER OF PAGES

20

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT UL

Methodology for an Operationally-Based Test Length Decision

by

Donald P. Gaver

Patricia A. Jacobs

Department of Operations Research
Naval Postgraduate School
Monterey, CA 93943
dgaver@wposmtp.nps.navy.mil

Abstract

Weapon systems that function destructively (e.g. missile or torpedo) are to be acquired in a lot of size m. Acceptance of the lot is based on the result of an operational test, administered to part of the lot: if the test results indicate positive operational value the lot is accepted and the remaining part of the lot if fielded; otherwise the lot is "rejected".

A test plan is designed that establishes an optimal number of weapon copies to test, given (models of) the operational gain of the fielded weapon under two tactical options, and the uncertainty in the weapon's predicted probability of success after the test is complete. The major test objective is to realize possible operational utility from the lot of items, and secondarily to demonstrate arbitrary levels of certainty.

Methodology for an Operationally-Based Test Length Decision

Donald P. Gaver

Patricia A. Jacobs

1. Problem Formulation

A new, or upgraded, item (for example, but not necessarily, a weapon) is to be produced and introduced into service provided it passes a test of its effectiveness and suitability. The test envisioned consists of employing the item in a typical mission and assessing the outcome as a success or failure. Tests are destructive: each test means that the tested item is lost, but its record is available: summarized as either success or failure. It is recognized that a more sensitive measure that accounts for *degree* of success, such as miss distance, may be available and preferable, but treatment of this option is deferred for the present. Likewise, non-destructive and reliability-growth settings can receive similar treatment; for discussion of a reliability growth problem in a similar context, see Gaver and Jacobs (1997).

Suppose that m items will be produced or bought in a lot. An option is to test t $(0 \le t \le m)$ of them and choose whether to accept or reject the entire lot on the basis of the test. Clearly a small test (e.g. based on firing a few missiles, perhaps 2) conveys less information about the item than does a larger one. But a large test means that only a few are left for field deployment. We intend to provide a formulation that links the *information* provided by a test to the *effect* of that

information upon eventual operational performance. We indicate that it is also possible to affect the decision as to how the item is "best" employed in the field, i.e. to assess the value based on simple tactical changes. The present formulation is intentionally simplified, but it forces a choice that is often only implicit in practice. The general methodology can be extended to include many more realistic details. We plan to do so in subsequent work.

2. Uncertainty, Information, and Operational Gain

Assume that the item has a constant but unknown probability of success, p. Whether to field a weapon depends on the value of p, which will ultimately be revealed, but still with uncertainty, with the assistance of a physical operational test. An option is to make t (t either 0 (no tests), 1, 2, ...m) tests and then decide whether p seems to be large enough. (Note that if $t = m$ maximum information will be procured, but nothing is left for use!)

Whether p is large enough can be based on a model of the weapon effectiveness as used operationally. The following is a simple option: when the weapon is actually used it is fired at an opponent; the latter, if missed, has the option to fire back. Letting v_w be the value of a *win*, meaning a weapon kill of the enemy target, and v_ℓ being the value of a loss, possibly meaning the loss of the weapon-firing platform, and q being the success probability of enemy counterfire (assumed known for simplicity), then

$$G_1(p,t;m) = (m-t)\left[v_w p - v_\ell(1-p)q\right] \tag{2.1}$$

is the total expected gain from fielding a weapon with success probability p having made t tests, so $(m - t)$ future engagements are actually possible. One can also account for the cost, and/or availability, of the actual missile; this is not done here.

2

The subscript on G_1 signifies that 1 weapon firing will occur per engagement. Alternatively we can contemplate firing a 2-weapon salvo, in which case we put

$$G_2(p,t;m) = \left\lfloor \frac{m-t}{2} \right\rfloor \left[v_w \left(1 - (1-p)^2 \right) - v_\ell (1-p)^2 q \right] \qquad (2.2)$$

where $\lfloor x \rfloor$ is the largest integer less than or equal to x. Another option is shoot-look-shoot, and there are many others; these are not considered in this paper.

The above, of course, is conditional on the particular operational setup envisioned, perhaps importantly characterized by range of engagement, but also target type, relative orientation, altitude, and other factors. Certainly both (unknown) p, and response probability, q, may depend sensitively on such conditions. In what follows we assume that all tests are conducted under one set of operational conditions and target tactics.

There are many different and more elaborate options for expression of net gain, but analysis of the above is itself instructive. A basic issue is, of course, how to specify numerical values for the gain and loss parameters, v_w and v_ℓ, or at least their ratio. Most literally, v_w could be equated to the monetary cost to an opponent of replacement of the item (e.g. targeted platform) that the missile attempts to kill; v_ℓ represents the same to the missile shooter. More cogently, the value chosen for v_w should reflect the military value of destroying the targeted opponent, this being very large if, for example, the opponent platform threatens a major component of one's own force, such as a battlegroup's carrier. The degree of that threat could be quantitatively assessed by exercise of an appropriate wargame or modeling exercise, conducted offline; such an exercise could also lead to an appropriate (relative) value for v_ℓ. Initially, however, expert judgment would likely be used to set tentative values, thus supporting a test decision. The inevitable initial subjectivity should stimulate more intensive examination of tradeoff issues, and lead to better approaches to quantifying v_w

and v_ℓ, such as by modeling and simulation combined with careful scrutiny of historical information.

Decision Logic

If the decision maker *knows* the value of p then he opts not to test ($t = 0$) and presumably evaluates $G_i(p, 0; m)$: if it is (sufficiently) *positive* he *accepts* the system and fields it, while if it is *negative* he *rejects* the system, achieving a gain of zero. This is, of course, again oversimplified. For one thing, the particular weapon system's mission might be alternatively performed by a similar predecessor system, so the gain of the prospective new item should be compared to that of the predecessor. There might also be a quite different alternative to which it could be compared. Thus the gain functions above are merely illustrative, intended to be provocative and to stimulate analytical thinking (often the most important objective of a model). In any case the decision as to whether to accept or reject the system often reduces to assessing the evidence that the actual probability of success, p, exceeds some threshold value, $\underline{p}$. For instance for gain function G_1 it can be seen that in order for the acceptance system to have positive gain $\underline{p} = v_\ell q/(v_w + v_\ell q)$ is that threshold. The threshold picked is thus dictated by the actual operational situation envisioned, including opponent response and relative platform values.

Uncertainty in p: Bayes Approach to Acceptance, Given a Test

The decision to test t leads to acquisition of data assumed entirely summarized as *s successes* ($s = 0, 1, 2, \ldots, t$). These data can now be used to create a likelihood for p by use of a binomial model, and, if p is endowed with a beta prior, a beta posterior for p:

$$\pi(p; s, t) = B(\alpha', \beta')p^{\alpha'-1}(1-p)^{\beta'-1} \tag{2.3}$$

4

where $\alpha' = \alpha + s$, and $\beta' = \beta + t - s$, the values (α, β) characterize the original prior density for p. This is the classical conjugate prior setup, see Berger (1985), and is invoked for convenient and flexible illustration; other options could be used. Then a decision maker in possession of (2.3) should use it to evaluate the gain, which can be done by computing its expectation with respect to the posterior probability distribution. In the case of G_1 the linear form gives

$$E\big[G_1(p,t;m)\big|s,t\big] = (m-t)\big[v_w E[p|s,t] - v_\ell(1 - E[p|s,t])q\big] \tag{2.4}$$

for the expected gain if the system is fielded. An appropriate decision rule may be:

Field the system if $E\big[G_1(p,t;m)\big|s,t\big]$ is positive; otherwise "reject". This is equivalent to fielding if $E[p|s,t] > \underline{p} = v_\ell q/(v_w + v_\ell q)$. Because of the form of G_1 this is also equivalent to $s \geq \underline{s}(t)$.

How Much to Test

The previous step indicates what decision to make, *given* test number t, and outcome $s(t)$ successes. Now take the position of the decision maker *before* any tests are made. She must consider testing to any level, i.e. $t = 0, 1, 2, ..., m$. The prediction used must depend on the binomial model and upon the prior, which is assumed to be the same beta prior with parameters α and β

$$\pi(p) = B(\alpha, \beta)p^{\alpha-1}(1-p)^{\beta-1} \tag{2.5}$$

conditional upon p,

$$P\{s(t) = s|p,t\} = \binom{t}{s}p^s(1-p)^{t-s} \tag{2.6}$$

In order to predict $s(t)$ simply remove the condition on p:

5

$$b(s,t) \equiv P\{s(t) = s|t\} = \int_0^1 \binom{t}{s} p^s (1-p)^{t-s(t)} B(\alpha,\beta) p^{\alpha-1} (1-p)^{\beta-1} dp$$

$$= \binom{t}{s} \cdot \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \cdot \frac{\Gamma(\alpha+s)\Gamma(\beta+t-s)}{\Gamma(\alpha+\beta+t)} \qquad (2.7)$$

$$= b(s-1,t)\left[\left(\frac{t+1-s}{s}\right)\left(\frac{\alpha-1+s}{\beta+t-s}\right)\right].$$

The expectation of gain, i.e. $E[G_1]$ depends on $s(t)$, which is unknown at the time of test planning, but whose predictive probability distribution is (2.7), the well-known beta-binomial. Consequently predict the expectation of future gain by calculating

$$E\left\{E\left[G_1(p,t;m)|s(t),t\right]\right\} = (m-t)\left[(v_w + v_\ell q)\sum_{s \geq s(t)} \frac{\alpha+s}{\alpha+\beta+t} b(s,t) - v_\ell q \sum_{s \geq s(t)} b(s,t)\right]. \qquad (2.8)$$

This can in principle be evaluated for "all" t values, 0, 1, 2, ...; practically, start small and continue while the predicted expected gain first increases and then decreases – if it indeed starts small, increases, and then ultimately declines, as will often be the case. The optimal test time, t^*, can then be easily selected. It may well happen that an initial prior will be distributed optimistically, suggesting the decision not to test at all. Prudence suggests that such a decision be over-ridden, if only to try to reveal some totally unanticipated, and unmodeled, system flaw. Conversely, a test *might* suggest virtue in a system evaluated very poorly *a priori.* It is difficult to imagine that any serious decision maker would accept a new system without direct evidence as to its operability.

3. Acceptance Risk

In the previous section a procedure to determine a "best" number of tests, t^*, based on an operationally relevant gain was discussed. In this section we consider measures of *risk* associated with acceptance after t^* such tests. Recall

that acceptance means that $s(t^*) \geq \underline{s}(t^*)$ where $\underline{s}(t^*)$ is the smallest number of successes allowed in the t^* tests if acceptance is to occur. The measures of risk are as follows:

a) **The conditional expected value of the probability of future mission success, given Acceptance**

$$\bar{p}(t) = E[p|\text{Accept}] = \frac{\sum_{s \geq \underline{s}(t^*)} \left(\frac{\alpha + s}{\alpha + \beta + t^*}\right) b(s, t^*)}{\sum_{s \geq \underline{s}(t^*)} b(s, t^*)} \tag{3.1}$$

where $b(s, t)$ is defined in (2.7).

b) **The conditional distribution of the number of successes in the remaining $m - t^*$ weapons, given Acceptance**

Let M be the number of successes using the remaining $(m - t^*)$ weapons.

$$P\left\{M = k \middle| s(t^*) = s, \text{Accept}\right\}$$

$$= \int \binom{m - t^*}{k} p^k (1-p)^{m - t^* - k} \frac{\Gamma(\alpha + \beta + t^*)}{\Gamma(\alpha + s)\Gamma(\beta + t^* - s)}$$

$$\times p^{\alpha + s - 1}(1 - p)^{\beta + (t^* - s) - 1} dp \tag{3.2}$$

$$= \binom{m - t^*}{k} \frac{\Gamma(\alpha + \beta + t^*)}{\Gamma(\alpha + s)\Gamma(\beta + t^* - s)} \frac{\Gamma(\alpha + s + k)\Gamma(\beta - s + m - k)}{\Gamma(\alpha + \beta + m)}$$

for $s \geq \underline{s}(t^*)$.

Thus,

$$P\{M = k|\text{Accept}\} = \frac{\sum_{s \geq \underline{s}(t^*)} P\left\{M = k \middle| s(t^*) = s\right\} b(s, t^*)}{\sum_{s \geq \underline{s}(t^*)} b(s, t^*)} \tag{3.3}$$

7

where $b(s, t)$ is defined in (2.7). It is also instructive to compute the above given that the decision rule is overridden: the number of successes fell below $\underline{s}(t^*)$ but the system is accepted despite this fact.

4. A Two-Weapon Salvo is Fired at a Target

Once again assume that the weapon has a constant unknown probability of success, p. However, a salvo of two weapons is fired at a target. Presumably, this doctrine will allow acceptance of a weapon with a smaller value of p than if only one weapon were fired per target. The total expected gain from fielding a weapon with success probability p having made t tests, so $\left\lfloor \dfrac{(m-t)}{2} \right\rfloor$ future engagements are actually possible, is given by $G_2(p, t; m)$ in (2.2) where $\lfloor x \rfloor$ is the largest integer less than or equal to x. Now, however, individual missions are twice as costly as before in terms of weapon expenditure.

We assume in what follows that if t weapons are tested, then they are fired one at a time (this may well be poor testing practice). Thus the result of a test that fires t weapons is summarized as the number of successes to occur, $s(t)$, ($s(t) = 0, 1, \ldots, t$). In this case the expected gain if the system is fielded given $s(t)$ and t

$$
\begin{aligned}
E\big[G_2(p,t;m)\big|s,t\big] &= \left\lfloor \frac{m-t}{2} \right\rfloor \left[v_w E\Big[1-(1-p)^2\big|s,t\Big] - v_\ell\Big[E\big[(1-p)^2\big|s,t\big]q\Big]\right] \\
&= \left\lfloor \frac{m-t}{2} \right\rfloor \left[v_w - (v_w + v_\ell q)E\big[(1-p)^2\big|s,t\big]\right].
\end{aligned}
$$

(4.1)

Assuming the beta prior distribution, (2.5) as before, the predictive distribution of $s(t)$ is given by (2.7). The prediction of the expectation of future gain if test t weapons and accept the lot if the number of successes is greater than or equal to $\underline{s}(t)$ is

$$E\left\{E\left[G_2(p,t;m)\big|s(t),t\right]\right\}$$

$$= \sum_{s \geq \underline{s}(t)} \left\lfloor \frac{m-t}{2} \right\rfloor \left[v_w - (v_w + v_\ell q)\frac{(\beta+t-s+1)(\beta+t-s)}{(\alpha+\beta+t+1)(\alpha+\beta+t)} \right] b(s,t) \qquad (4.2)$$

A best number of tests, t_2^*, is that t-value, here t_2^* and $\underline{s}\left(t_2^*\right)$ that maximizes the

right-hand side of (4.2).

Acceptance Risk

In this section we consider measures of risk associated with acceptance after t^*

tests. Recall acceptance means $s(t^*) \geq \underline{s}(t^*)$ where $\underline{s}(t^*)$ is the smallest number of

successes allowed in the t^* tests for acceptance to occur.

a) **The conditional expected value of the probability of future mission success given Acceptance, when two weapons are salvoed per target**

$$\bar{p}_2(t) = E\left[1-(1-p)^2\big|\text{Accept}\right] = \frac{\displaystyle\sum_{s \geq \underline{s}(t)} \left[1 - \frac{(\beta+t-s+1)(\beta+t-s)}{(\alpha+\beta+t+1)(\alpha+\beta+t)} \right] b(s,t)}{\displaystyle\sum_{s \geq \underline{s}(t)} b(s,t)} \qquad (4.3)$$

where $b(s,t)$ is defined in (2.7).

b) **The conditional distribution of the number of successful target kills in the remaining $m - t^*$ weapons, given Acceptance, when two weapons are used per target**

Let M be the number of successes using the remaining $(m - t^*)$ weapons. Since

two weapons are used per target, $N(m, t^*) = \left\lfloor \dfrac{m-t^*}{2} \right\rfloor$ targets can be engaged.

$$P\{M = k|s,\text{Accept}\} = \int \binom{N\left(m,t^*\right)}{k} \left[1-(1-p)^2\right]^k \left[(1-p)^2\right]^{N\left(m,t^*\right)-k}$$

$$\times \frac{\Gamma\left(\alpha+\beta+t^*\right)}{\Gamma(\alpha+s)\Gamma\left(\beta+t^*-s\right)} p^{\alpha+s-1}(1-p)^{\beta+\left(t^*-s\right)-1} dp \qquad (4.4,a)$$

9

$$= \binom{N(m,t^*)}{k} \sum_{n=0}^{k} \left\{ (-1)^n \binom{k}{n} \frac{\Gamma(\alpha + \beta + t^*)}{\Gamma(\alpha + s)\Gamma(\beta + t^* - s)} \right.$$

$$\left. \times \frac{\Gamma(\alpha + s)\Gamma(\beta + t^* - s + 2n + 2N(m,t^*) - 2k)}{\Gamma(\alpha + \beta + t^* + 2n + 2N(m,t^*) - 2k)} \right\}$$

(4.4,b)

for $s \geq \underline{s}(t^*)$.

Thus,

$$P\{M = k | \text{Accept}\} = \frac{\sum\limits_{s \geq \underline{s}(t^*)} P\{M = k | s\} b(s,t^*)}{\sum\limits_{s \geq \underline{s}(t^*)} b(s,t^*)}$$

where $b(s,t)$ is defined in (2.7). Under certain circumstances, e.g. for large m-values, the alternating series form (4.4,b) becomes ill-conditioned, and it is preferable to carry out a numerical integration to evaluate (4.4,a).

5. Numerical Examples

We now exhibit numerical examples to illustrate the theory. Tables 1 – 3 present results of evaluating best policies and their acceptance risks for both one-weapon and two-weapon salvoes. In all the examples the beta prior has mean $\alpha / \alpha + \beta = 5/6 = 0.83$ for $\alpha = 5$ and $\beta = 1$. Table 1 presents results for the one-weapon salvo for number of missiles in the (small) lot sizes $m = 15$ and 30; value of win, $v_w = 1$; success probability of enemy counterfire $q = 0.7$; and various values of loss, v_ℓ. Table 2 presents corresponding results for a two-weapon salvo. It is apparent that as the value of loss becomes higher (i.e. at $v_\ell = 25$) the optimal number of tests suggested, and that must be passed, explodes for $m = 15$, ($t^* = 14$, leaving one(!) to be fielded), and is over half the lot size for m = 30; this in spite of a strong prior probability of single shot success. This requirement is much

reduced if two-weapon salvoes are fired, with expected gain much higher as well. But the two-missile salvoes only apply to one-half as many missions.

For the large missile lot of $m = 100$ it is again apparent that the possible loss of a highly expensive platform ($v_\ell = 25$, vs. $v_w = 1$) implies that conservative testing be done: in this case fielding is recommended if just over one-third of the lot is tested with, literally, no more than one failure allowed. But our methods allow estimation of expected gain whatever the test outcome, usefully informing a decision maker of the (conditional expected) gain whatever the test outcomes. This is more realistic in practice than is adherence to a "drop-dead" binary policy.

There is apparently a premium on buying, and testing, relatively large missile lots. The unmodeled costs include that of the possibility that the opponent will adapt to or counter new designs or tactics (e.g. raise q, or change engagement conditions) before the lot is consumed. It is also possible that a new system design will render the current design items obsolete or degraded in storage before all are used up; however, an upgrade may be possible.

TABLE 1

No Testing
(Prior-Based)

$\alpha = 5, \ \beta = 1, \ v_w = 1, \ q = 0.7$

One-Weapon Salvo

m	Expected Prob. of Success	Prob. of All Successes	Expected # of Successes (Std. Dev.)
15	0.83	0.25	12.5 (2.5)
25	0.83	0.14	25 (4.6)

Testing: Optimal Number of Tests and Risks

$\alpha = 5, \ \beta = 1, \ v_w = 1, \ q = 0.7$

One-Weapon Salvo

v_ℓ	m	Optimal # of tests	Min. # of Successes Needed to Accept	Max. Gain	Conditional Expected Prob. of Success Given Accept	Conditional Prob. of All Successes in Remaining Missiles Given Accept	Expected Number of Successes in Remaining Missiles (Std. Dev.)
5	15	1	1	4.2	0.86	0.3	12 (2.12)
15	15	9	9	0.5	0.93	0.7	5.6 (0.70)
25	15	14	14	0.02	0.95	0.95	0.95 (0.05)
5	30	2	2	8.8	0.86	0.2	24.5 (3.5)
15	30	11	11	1.92	0.94	0.46	17.9 (1.45)
25	30	18	18	0.60	0.96	0.66	11.5 (0.83)

TABLE 2
No Testing
(Prior-Based)
$\alpha = 5$, $\beta = 1$, $v_w = 1$, $q = 0.7$
Two-Weapon Salvo

m	Expected Prob. of Success	Prob. of All Successes	Expected # of Successes (Std. Dev.)
15	0.95	0.78	6.67 (0.75)
25	0.95	0.66	14.3 (1.4)

Testing: Optimal Number of Tests and Risks
$\alpha = 5$, $\beta = 1$, $v_w = 1$, $q = 0.7$
Two-Weapon Salvo

v_ℓ	m	Optimal # of tests	Min. # of Successes Needed to Accept	Max. Gain	Conditional Expected Prob. of Success Given Accept	Conditional Prob. of All Successes in Remaining Missiles Given Accept	Expected Number of Successes in Remaining Missiles (Std. Dev.)
5	15	0	0	5.5	0.95	0.78	6.67 (0.75)
15	15	1	1	3.4	0.96	0.82	6.75 (0.62)
25	15	3	3	2.2	0.98	0.89	5.87 (0.42)
5	30	0	0	11.0	0.95	0.66	14.3 (1.4)
15	30	4	3	7.08	0.97	0.76	12.6 (0.86)
25	30	6	5	5.10	0.98	0.81	11.7 (0.66)

TABLE 3

Testing: Optimal Number of Tests and Risks

$$\alpha = 5, \quad \beta = 1, \quad v_w = 1, \quad q = 0.7$$

One-Weapon Salvo

v_ℓ	m	Optimal # of tests	Min. # of Successes Needed to Accept	Max. Gain	Conditional Expected Prob. of Success Given Accept	Mode of Conditional Dist. of # of Successes in Remaining Missiles	Expected Number of Successes in Remaining Missiles (Std. Dev.)
0	100	0	0	83.3			
5	100	10	8	33.3	0.89	Mode = 90 0.07	73.6 (8.6)
15	100	23	22	10.7	0.95	Mode = 77 0.18	73.1 (3.8)
25	100	37	36	5.2	0.97	Mode = 63 0.29	60.9 (2.4)
No testing	100	0	0		0.83	Mode = 100 0.05	83.3 (14.5)

TABLE 4

Two-Weapon Salvo

v_ℓ	m	Optimal # of tests	Min. # of Successes Needed to Accept	Max. Gain	Conditional Expected Prob. of Success Given Accept	Mode of Conditional Dist. of # of Successes in Remaining Missiles	Expected Number of Successes in Remaining Missiles (Std. Dev.)
5	100	0	0	39.3	0.95	Mode = 50 0.46	47.6 (4.0)
15	100	8	6	26.3	0.97	Mode = 46 0.57	44.8 (2.1)
25	100	10	8	20.7	0.98	Mode = 45 0.61	44.1 (1.7)

References

Berger, J.O., *Statistical Decision Theory and Bayesian Analysis*, Second Edition, Springer-Verlag, New York, 1985.

Gaver, D.P. and Jacobs, P.A., (1997). "Testing or fault-finding for reliability growth: a missile destructive-test example," *Naval Research Logistics* (to appear).

DISTRIBUTION LIST

1. Research Office (Code 09) 1
    Naval Postgraduate School
    Monterey, CA 93943-5000

2. Dudley Knox Library (Code 013) 2
    Naval Postgraduate School
    Monterey, CA 93943-5002

3. Defense Technical Information Center..2
    8725 John J. Kingman Rd., STE 0944
    Ft. Belvoir, VA 22060-6218

4. Therese Bilodeau...1
    Dept of Operations Research
    Naval Postgraduate School
    Monterey, CA 93943-5000

5. Prof. Donald P. Gaver (Code OR/Gv)...5
    Naval Postgraduate School
    Monterey, CA 93943-5000

6. Prof. Patricia A. Jacobs (Code OR/Jc) 5
    Naval Postgraduate School
    Monterey, CA 93943-5000

7. Dr. J. Abrahams...1
    Code 111, Room 607
    Mathematical Sciences Division, Office of Naval Research
    800 North Quincy Street
    Arlington, VA 22217-5000

8. Prof. D. R. Barr 1
    Dept. of Systems Engineering
    U.S. Military Academy
    West Point, NY 10996

9. Center for Naval Analyses 1
    4401 Ford Avenue
    Alexandria, VA 22302-0268

10. Prof. H. Chernoff...1
    Department of Statistics
    Harvard University
    1 Oxford Street
    Cambridge, MA 02138